

We Claim:

1. A method of analyzing expressed opinions comprising the steps of:
 - 5 parsing words of at least one text-based document as parts of speech;
 - extracting regular expressions from the document by matching at least one regular expression rule with the parsed parts of speech; and
 - 10 categorizing extracted regular expressions into representative categories of semantic orientation by analyzing the words comprising the extracted regular expressions.
2. The method of claim 1, wherein extracted opinions are characterised by the
15 polarity of their sentiment into at least positive and negative categories.
3. The method of claim 1, wherein the representative categories are identified as favourable, unfavourable, and indifferent.
- 20 4. The method of claim 1, further comprising the step of graphically displaying the relative proportions of the differently categorised expressions.
5. The method of claim 1, further comprising the step of storing a collection of regular expression rules indicative of expressed opinions.
- 25 6. The method of claim 1, further comprising the step of associating expressed opinions with topics to which the expressed opinions relate.
7. The method of claim 1, further comprising the step of accessing a natural
30 language database to determine the semantic orientation of words of a regular expression.

- 8.. The method of claim 1, further comprising the step of accessing a lexical reference to identify possible synonyms or antonyms for words of a regular expression.
- 5 9.. The method of claim 1, further comprising the step of determining a predominant connotation of an extracted regular expression by performing at least one of the following steps in relation to words of the regular expression:
- (i) determining the semantic orientation of the words;
 - 10 (ii) determining the semantic orientation of at least one of synonyms and antonyms for the words; and
 - (iii) determining the semantic orientation of morphological stems for the words.
- 15 10.. A computer program product comprising computer software recorded on a computer-readable medium for performing the steps of:
- parsing words of at least one text-based document as parts of speech;
 - 20 extracting regular expressions from the document by matching at least one regular expression rule with the parsed parts of speech; and
 - categorizing extracted regular expressions into representative categories of semantic orientation by analyzing the words comprising the extracted regular expressions.
- 25
- 11.. A computer system comprising computer software recorded on a computer-readable medium for performing the steps of:
- 30 parsing words of at least one text-based document as parts of speech;
 - extracting regular expressions from the document by matching at least one regular expression rule with the parsed parts of speech; and

categorizing extracted regular expressions into representative categories of semantic orientation by analyzing the words comprising the extracted regular expressions.